

---

# Supplementary: RCSTAT: A Statistical Framework for using Relative Contextualization in Transformers

---

Anonymous Author(s)

Affiliation

Address

email

Table 1: **Value Error Rate (VER)** on the QMSum dataset across different compression ratios (50%, 60%, 70%) for LLaMA-3.2-3B and LLaMA-3.1-8B Instruct models. **RCStat (IOT)** assumes *Independent Output Tokens*, while **RCStat (Non-IOT)** does not assume any independence. Here, lower is better.

Model	Method	50%	60%	70%
3B	TOVA	0.2408	0.3103	0.3905
	RCSTAT (IOT)	0.1956	0.2648	0.3402
	RCSTAT (Non-IID)	<b>0.1571</b>	<b>0.2295</b>	<b>0.3066</b>
8B	TOVA	0.2177	0.2859	0.3639
	RCSTAT (IOT)	0.1615	0.2290	0.3007
	RCSTAT (Non-IID)	<b>0.1043</b>	<b>0.2034</b>	<b>0.2836</b>

## 1 Relative Contextualization distribution: Head-level Analysis

We show the per-head distribution of relative contextualization, in terms of the upper bound overlap area, in Figures 1, 2, and 3 for QMSum, Squadv2, and 2WikiMultiHop datasets, respectively. The percentile values of these distributions are shown in Figures 4, 5, and 6 respectively. These figures provide empirical evidence of our statement in the conclusion section: “the most influential contextualization heads consistently reside in the model’s middle layers, corroborating prior findings.” This can also be observed in Figures 4c and 4d in the main paper, where the high-scoring heads correspond to the middle layers: layer indices are shown in the x-axis labels.

## 2 Complete Experimental Statistics for KV-compression results

Please find the results of Value Error Rate (VER) inside the VER folder. For the baseline methods, the mean and standard deviations of VER for different compression ratios are saved in csv files with the naming format <dataset>\_<model>\_baseline\_df.csv, where the dataset field can be 2WikiMultiHop, QMSum, or SquAD, and the model field can be 3b or 8b. Similarly, the mean and standard deviations of VER and the mean and standard deviations of the compression ratios for different threshold multipliers are saved in csv files with the naming format <dataset>\_<model>\_proposed\_df.csv. Similarly, the results for Rouge1 and RougeL can be found in the All\_Rouges folder.

### 2.1 Independence assumption of generated tokens

The result in Table 1 shows that the fidelity of value vectors is higher when RCSTAT is executed without assuming independence for the random variables corresponding to  $\langle q, k \rangle$  of generated tokens.

Table 2: Additional result for the attribution task, when head selection is based on Relative contextualization applied on post-softmax.

Model	QuoteSum	VERI-GRAN
GPT-3.5 (inline)	90.18	26.40
GPT-4 (inline)	90.59	62.11
BM25	75.72	68.20
GTR	72.57	53.15
MT5	89.24	67.43
LLaMA-7B (HS)	87.51	77.33
Mistral-7B (HS)	89.95	77.71
L3.1-8B (all heads)	90.54	77.91
L3.1-8B (least RC post-softmax, $k=20$ )	<b>35.72</b>	<b>4.69</b>
L3.1-8B (least RC pre-softmax, $k=20$ )	29.49	2.81
L3.1-8B (most RC post-softmax, $k=20$ )	90.03	71.25
L3.1-8B (most RC pre-softmax, $k=20$ )	<b>93.91</b>	<b>79.37</b>

Nonetheless, even with the independence assumption, RCSTAT outperforms TOVA, which is the best-performing method in our experiments for the main paper.

### 3 Additional Results for Attribution Experiments

#### 3.1 Qualitative Comparison of Attribution Strategies at Layer 15

To better understand the effectiveness of various attention-based attribution strategies, we compare three different approaches using attention maps from Layer 15 of our model: (1) the mean attention across all heads, (2) the top-scoring head according to our attribution scoring technique, and (3) the worst-scoring head by the same measure. All methods were evaluated on the same input setup: a sales report document with the question “*What were the product sales on November 21st?*” and the answer “*The product sales on November 21st were \$177.00.*”

Figure 7 presents the attention heatmaps produced by each of the three strategies. The top-scoring head (Head 30, Figure 7a) yields a sharply focused attribution map, precisely attending to tokens corresponding to the correct numerical value. In contrast, the mean attention across all heads (Figure 7b) produces a reasonable heatmap but also attends to several unrelated tokens, leading to less interpretable attributions. Finally, the worst-scoring head (Head 16, Figure 7c) demonstrates diffuse and uninformative attention, highlighting mostly irrelevant tokens.

These observations qualitatively validate our scoring technique for identifying high-quality attribution heads and demonstrate that selectively using the best attention heads can significantly improve interpretability.

#### 3.2 Quantitative Comparison for Post-softmax and pre-softmax

In Table 2 we present the quantitative result of what happens if the head selection is based on the RC applied to post-softmax attention weights instead of pre-softmax attention logits. We observe that, when top heads are based on post-softmax weights, the attribution accuracy is lower than that of pre-softmax, whereas when the bottom heads are based on post-softmax weights, the attribution accuracy is not lower compared to pre-softmax. This shows that the information that can distinguish head-level importance is present more in pre-softmax logit values than in post-softmax weights.

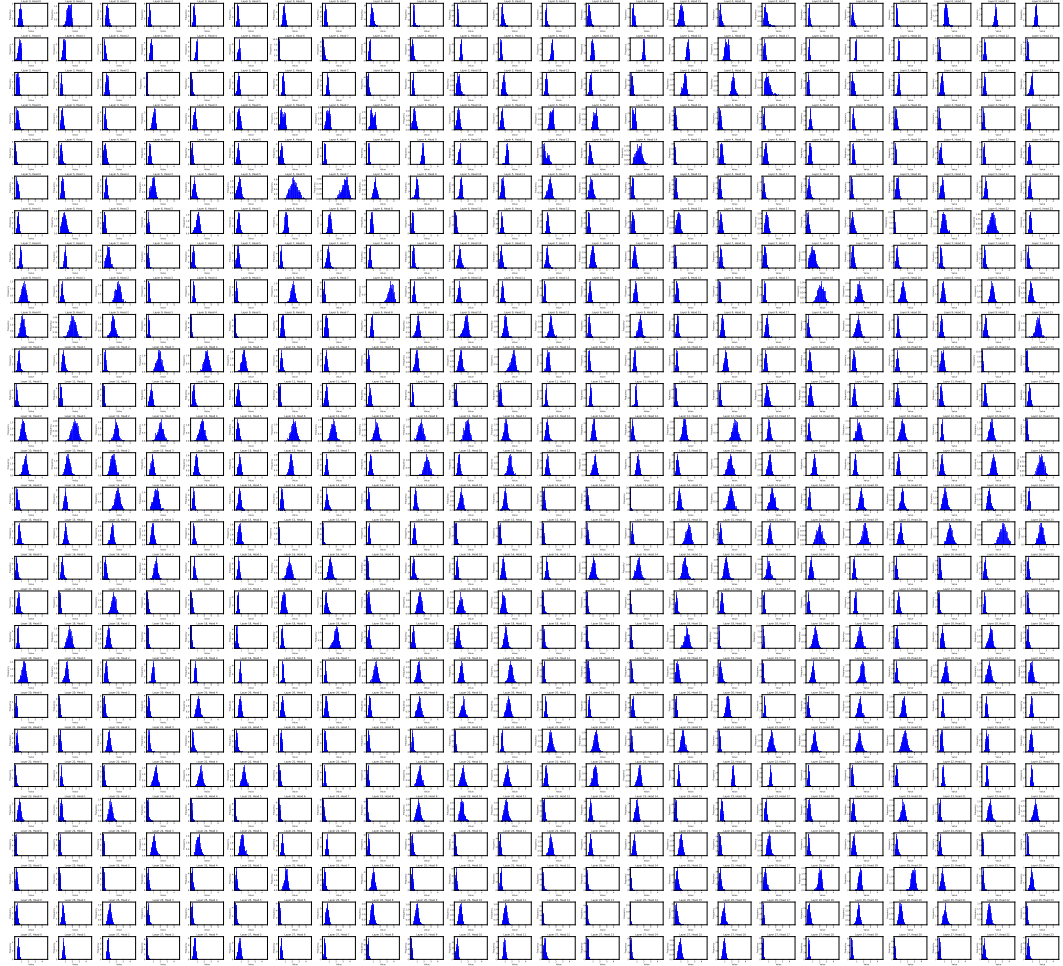


Figure 1: (see on screen) The distribution of RC upper bound (overlap area) for QmSum dataset. The first (last) row corresponds to heads in the first (last) layer.

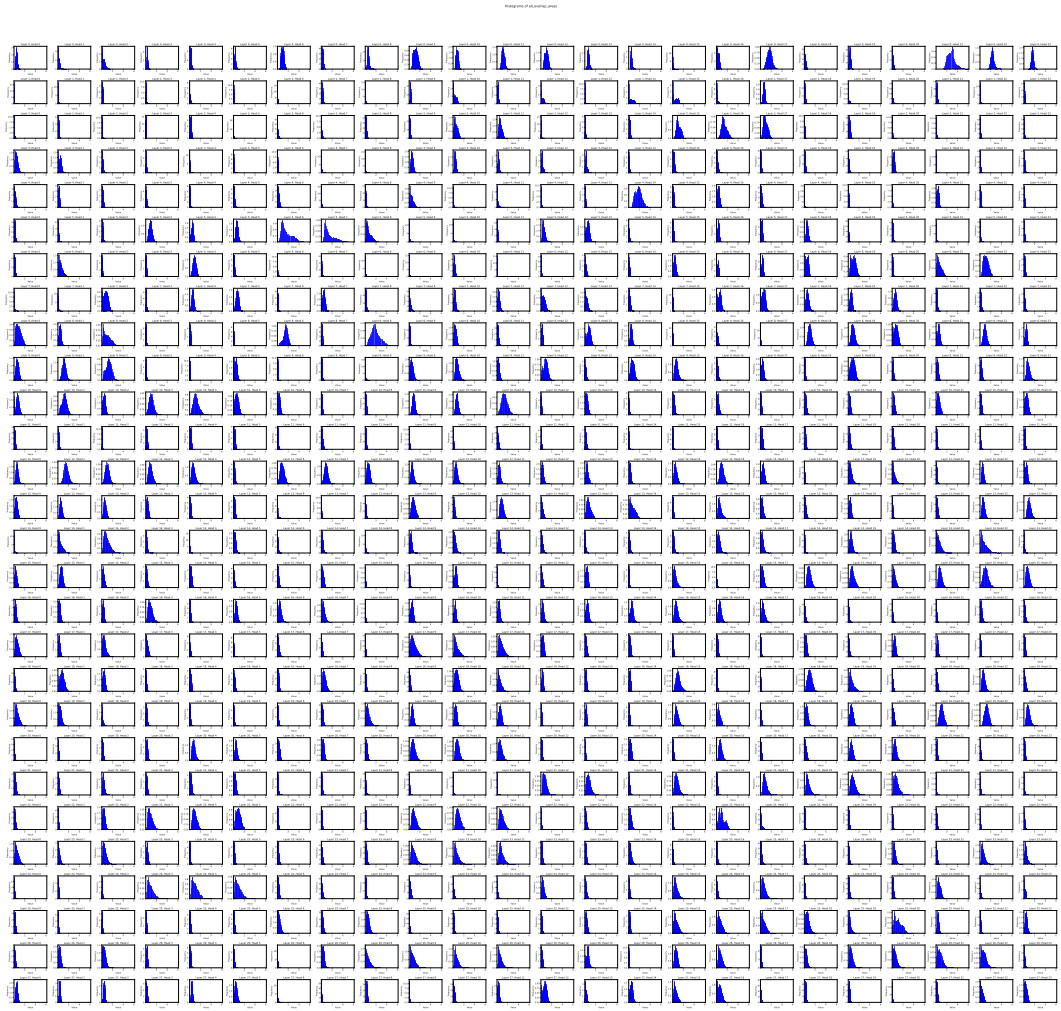


Figure 2: (see on screen) The distribution of RC upper bound (overlap area) for Squad v2 dataset. The first (last) row corresponds to heads in the first (last) layer.

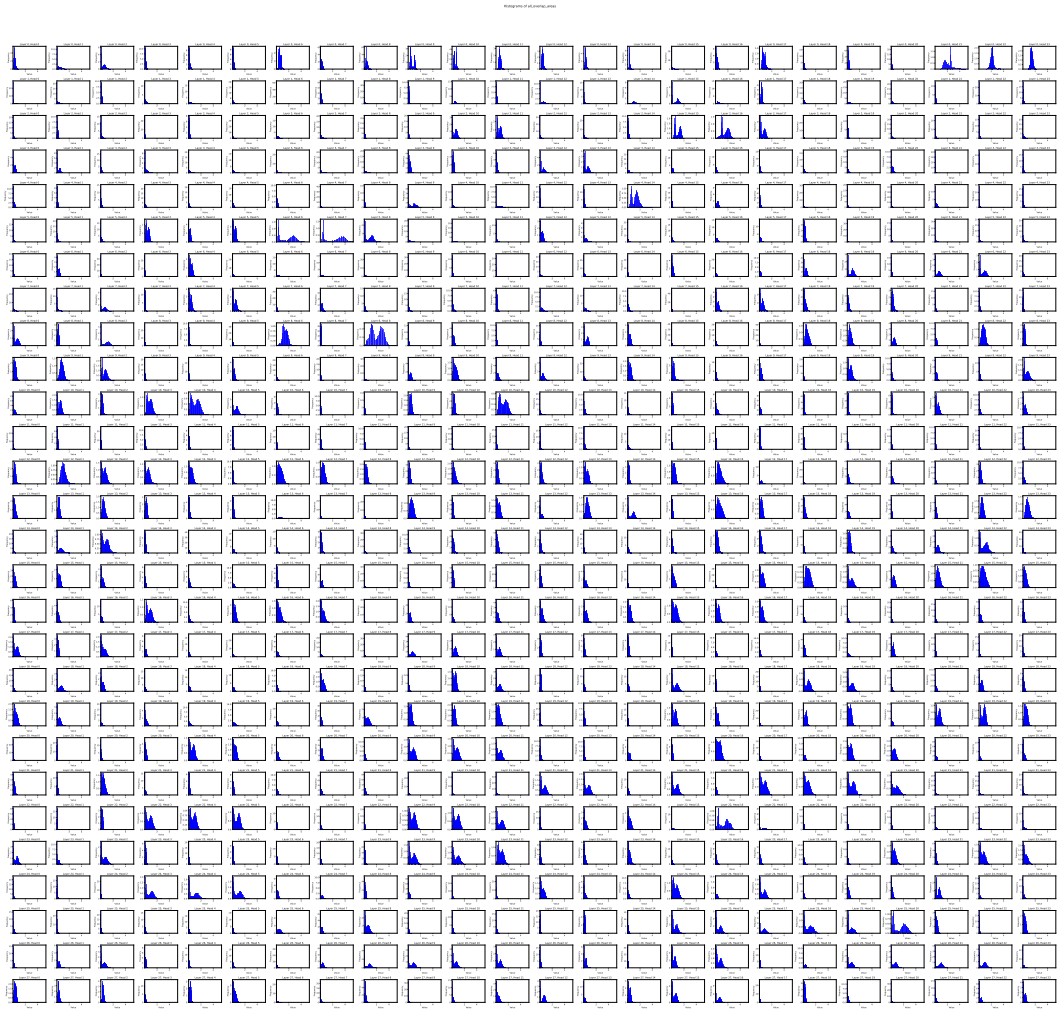


Figure 3: (see on screen) The distribution of RC upper bound (overlap area) for 2WikiMultiHop dataset. The first (last) row corresponds to heads in the first (last) layer.

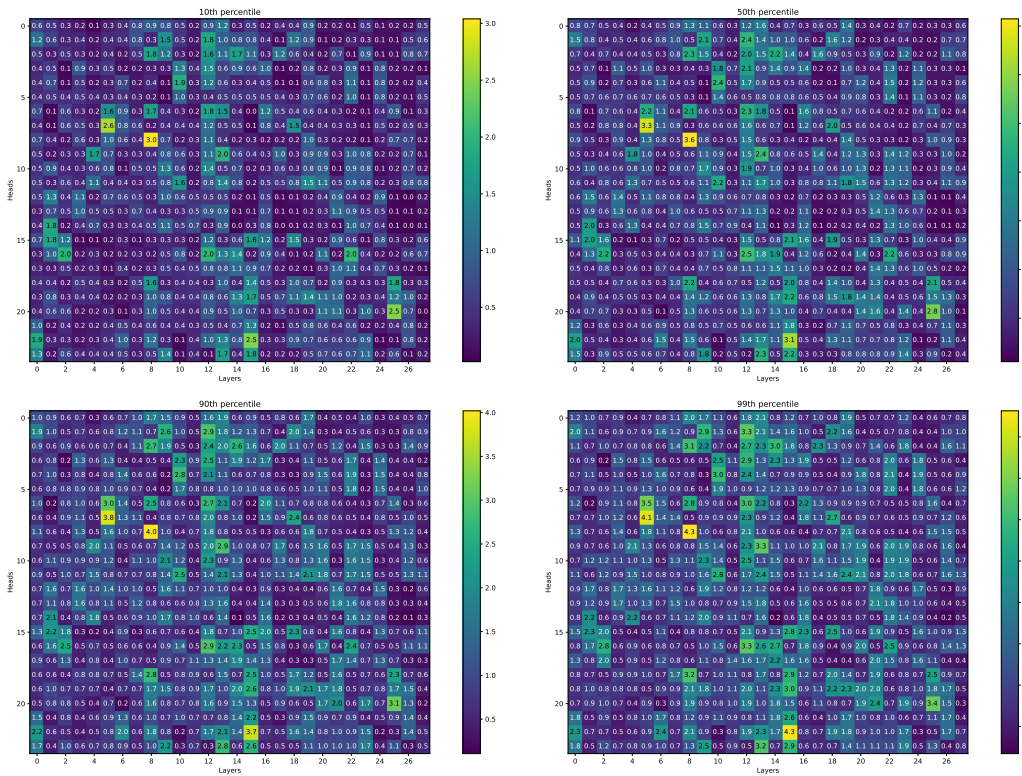


Figure 4: (see on screen) Percentiles of RC upper bound (overlap area) across the QmSum dataset

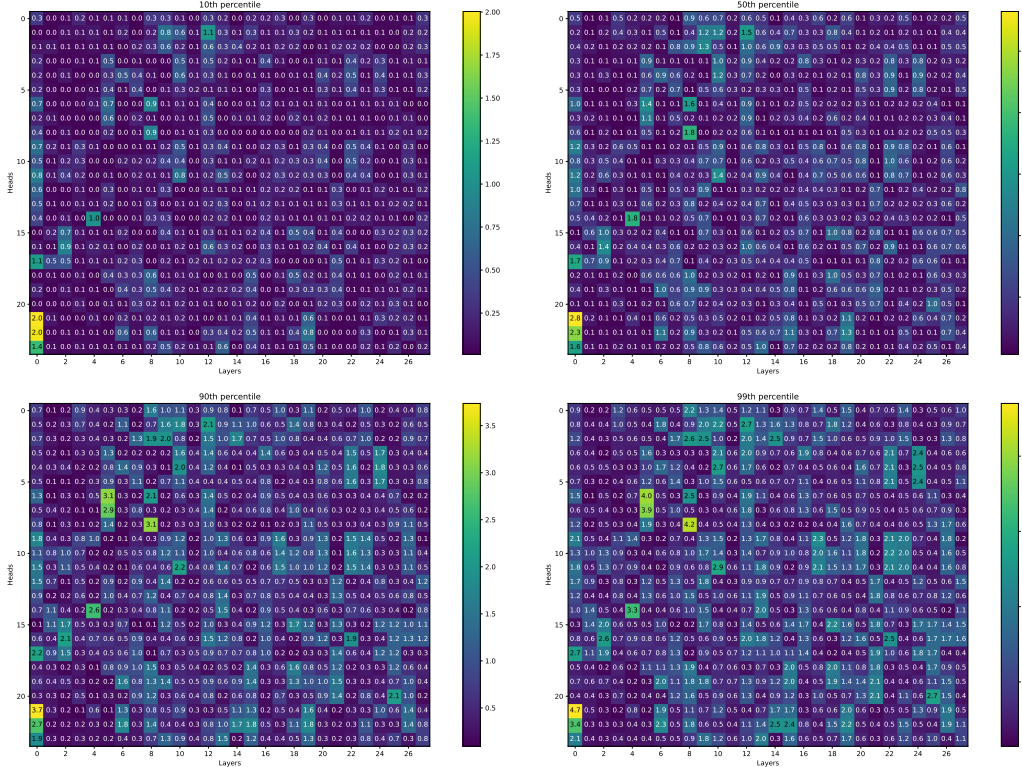


Figure 5: (see on screen) Percentiles of RC upper bound (overlap area) across the Squad dataset

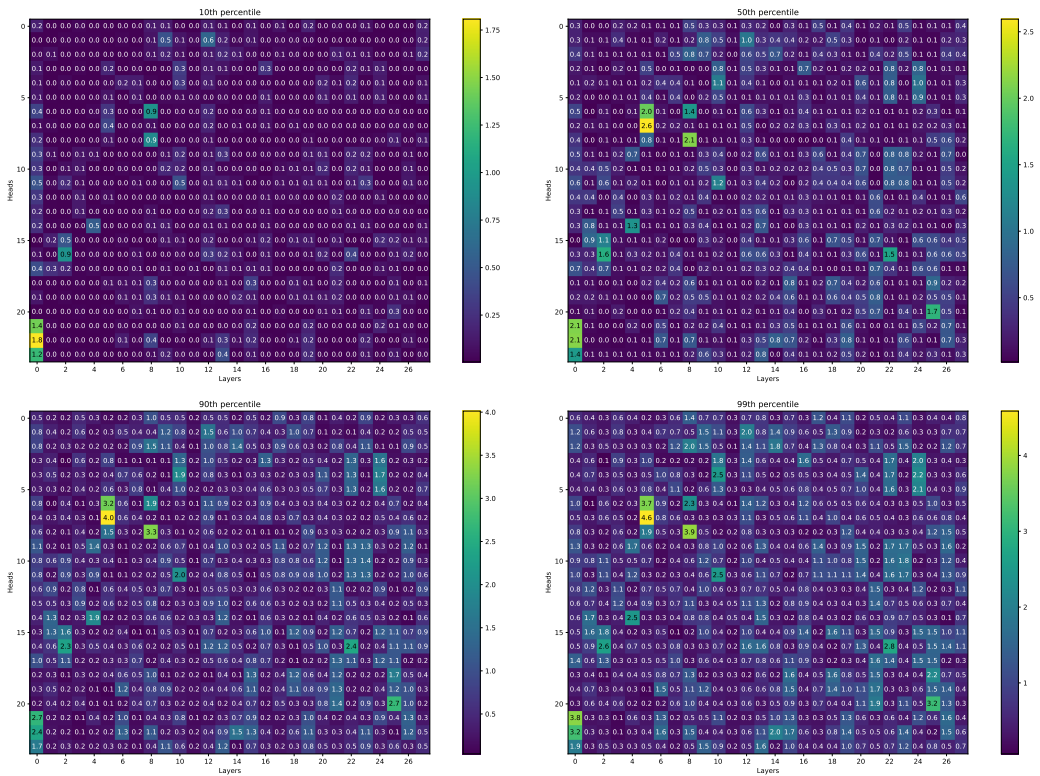


Figure 6: (see on screen) Percentiles of RC upper bound (overlap area) across the 2WikiMultiHop dataset



SALES SUMMARY REPORT

Location(s): Wax It Skin Studio  
Period: November 1 - November 30, 2023

Date	# Sales	# Services	Service Sales	# Products	Product Sales	Net Total	Taxes	Tips	Gross Total
Nov 1	10	5	\$500.00	7	\$343.95	\$1,174.45	\$28.89	\$180.88	\$1,384.22
Nov 2	17	18	\$1,330.00	9	\$519.60	\$1,850.10	\$43.65	\$226.65	\$2,120.40
Nov 3	5	5	\$621.00	3	\$208.00	\$829.00	\$17.48	\$124.20	\$970.68
Nov 6	9	8	\$972.00	3	\$150.00	\$1,122.00	\$12.60	\$180.45	\$1,315.05
Nov 7	7	6	\$675.00	1	\$58.00	\$733.00	\$4.87	\$126.75	\$864.62
Nov 8	6	6	\$582.00	3	\$223.00	\$805.00	\$18.73	\$82.20	\$905.93
Nov 9	12	14	\$1,263.75	3	\$218.00	\$1,481.75	\$18.31	\$216.95	\$1,717.01
Nov 10	4	6	\$385.00	1	\$32.00	\$417.00	\$2.69	\$77.00	\$496.69
Nov 11	1	2	\$270.00	0	\$0.00	\$270.00	\$0.00	\$54.00	\$324.00
Nov 13	13	16	\$1,586.00	2	\$150.00	\$1,736.00	\$12.60	\$266.45	\$2,015.05
Nov 14	15	21	\$2,011.00	0	\$0.00	\$2,011.00	\$0.00	\$414.95	\$2,425.95
Nov 15	10	8	\$745.00	3	\$213.00	\$958.00	\$17.90	\$118.00	\$1,093.90
Nov 16	9	11	\$1,565.00	6	\$513.00	\$2,078.00	\$43.10	\$312.25	\$2,433.35
Nov 17	17	24	\$1,797.50	0	\$0.00	\$1,797.50	\$0.00	\$388.25	\$2,185.75
Nov 18	2	0	\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00	\$0.00
Nov 19	2	0	\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00	\$0.00
Nov 20	15	17	\$1,777.50	1	\$36.95	\$1,814.45	\$3.10	\$336.00	\$2,153.55
Nov 21	17	24	\$2,640.00	4	\$177.00	\$2,817.00	\$14.88	\$313.00	\$3,144.88
Nov 22	10	8	\$945.00	5	\$238.00	\$1,183.00	\$20.01	\$77.50	\$1,280.51
Nov 24	2	0	\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00	\$0.00

Page 1 of 2

SALES SUMMARY REPORT

Location(s): Wax It Skin Studio  
Period: November 1 - November 30, 2023

Date	# Sales	# Services	Service Sales	# Products	Product Sales	Net Total	Taxes	Tips	Gross Total
Nov 1	10	5	\$500.00	7	\$343.95	\$1,174.45	\$28.89	\$180.88	\$1,384.22
Nov 2	17	18	\$1,330.00	9	\$519.60	\$1,850.10	\$43.65	\$226.65	\$2,120.40
Nov 3	5	5	\$621.00	3	\$208.00	\$829.00	\$17.48	\$124.20	\$970.68
Nov 6	9	8	\$972.00	3	\$150.00	\$1,122.00	\$12.60	\$180.45	\$1,315.05
Nov 7	7	6	\$675.00	1	\$58.00	\$733.00	\$4.87	\$126.75	\$864.62
Nov 8	6	6	\$582.00	3	\$223.00	\$805.00	\$18.73	\$82.20	\$905.93
Nov 9	12	14	\$1,263.75	3	\$218.00	\$1,481.75	\$18.31	\$216.95	\$1,717.01
Nov 10	4	6	\$385.00	1	\$32.00	\$417.00	\$2.69	\$77.00	\$496.69
Nov 11	1	2	\$270.00	0	\$0.00	\$270.00	\$0.00	\$54.00	\$324.00
Nov 13	13	16	\$1,586.00	2	\$150.00	\$1,736.00	\$12.60	\$266.45	\$2,015.05
Nov 14	15	21	\$2,011.00	0	\$0.00	\$2,011.00	\$0.00	\$414.95	\$2,425.95
Nov 15	10	8	\$745.00	3	\$213.00	\$958.00	\$17.90	\$118.00	\$1,093.90
Nov 16	9	11	\$1,565.00	6	\$513.00	\$2,078.00	\$43.10	\$312.25	\$2,433.35
Nov 17	17	24	\$1,797.50	0	\$0.00	\$1,797.50	\$0.00	\$388.25	\$2,185.75
Nov 18	2	0	\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00	\$0.00
Nov 19	2	0	\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00	\$0.00
Nov 20	15	17	\$1,777.50	1	\$36.95	\$1,814.45	\$3.10	\$336.00	\$2,153.55
Nov 21	17	24	\$2,640.00	4	\$177.00	\$2,817.00	\$14.88	\$313.00	\$3,144.88
Nov 22	10	8	\$945.00	5	\$238.00	\$1,183.00	\$20.01	\$77.50	\$1,280.51
Nov 24	2	0	\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00	\$0.00

Page 1 of 2

Date	# Sales	# Services	Service Sales	# Products	Product Sales	Net Total	Taxes	Tips	Gross Total
Nov 25	3	0	\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00	\$0.00
Nov 26	2	0	\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00	\$0.00
Nov 27	6	0	\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00	\$0.00
Nov 28	16	19	\$2,565.00	2	\$123.00	\$2,708.00	\$10.34	\$490.25	\$3,208.59
Nov 29	10	13	\$1,270.00	0	\$0.00	\$1,270.00	\$0.00	\$284.90	\$1,554.90
Nov 30	22	18	\$1,685.50	12	\$720.80	\$2,406.30	\$60.56	\$293.00	\$2,759.86
Total	242	252	\$25,537.25	65	\$3,924.38	\$29,461.55	\$329.71	\$4,563.23	\$34,354.49

\* The service sales include the value of applied packages.

Date	# Sales	# Services	Service Sales	# Products	Product Sales	Net Total	Taxes	Tips	Gross Total
Nov 25	3	0	\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00	\$0.00
Nov 26	2	0	\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00	\$0.00
Nov 27	6	0	\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00	\$0.00
Nov 28	16	19	\$2,565.00	2	\$123.00	\$2,708.00	\$10.34	\$490.25	\$3,208.59
Nov 29	10	13	\$1,270.00	0	\$0.00	\$1,270.00	\$0.00	\$284.90	\$1,554.90
Nov 30	22	18	\$1,685.50	12	\$720.80	\$2,406.30	\$60.56	\$293.00	\$2,759.86
Total	242	252	\$25,537.25	65	\$3,924.38	\$29,461.55	\$329.71	\$4,563.23	\$34,354.49

\* The service sales include the value of applied packages.

(a) Top-scoring head (Head 30)

(b) Mean of all heads

SALES SUMMARY REPORT

Location(s): Wax It Skin Studio  
Period: November 1 - November 30, 2023

Date	# Sales	# Services	Service Sales	# Products	Product Sales	Net Total	Taxes	Tips	Gross Total
Nov 1	10	8	\$830.50	7	\$343.95	\$1,174.45	\$28.89	\$180.88	\$1,384.22
Nov 2	17	18	\$1,330.50	9	\$519.60	\$1,850.10	\$43.65	\$226.65	\$2,120.40
Nov 3	5	5	\$621.00	3	\$208.00	\$829.00	\$17.48	\$124.20	\$970.68
Nov 6	9	8	\$972.00	3	\$150.00	\$1,122.00	\$12.60	\$180.45	\$1,315.05
Nov 7	7	6	\$675.00	1	\$58.00	\$733.00	\$4.87	\$126.75	\$864.62
Nov 8	6	6	\$582.00	3	\$223.00	\$805.00	\$18.73	\$82.20	\$905.93
Nov 9	12	14	\$1,263.75	3	\$218.00	\$1,481.75	\$18.31	\$216.95	\$1,717.01
Nov 10	4	6	\$385.00	1	\$32.00	\$417.00	\$2.69	\$77.00	\$496.69
Nov 11	1	2	\$270.00	0	\$0.00	\$270.00	\$0.00	\$54.00	\$324.00
Nov 13	13	16	\$1,586.00	2	\$150.00	\$1,736.00	\$12.60	\$266.45	\$2,015.05
Nov 14	15	21	\$2,011.00	0	\$0.00	\$2,011.00	\$0.00	\$414.95	\$2,425.95
Nov 15	10	8	\$745.00	3	\$213.00	\$958.00	\$17.90	\$118.00	\$1,093.90
Nov 16	9	11	\$1,565.00	6	\$513.00	\$2,078.00	\$43.10	\$312.25	\$2,433.35
Nov 17	17	24	\$1,797.50	0	\$0.00	\$1,797.50	\$0.00	\$388.25	\$2,185.75
Nov 18	2	0	\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00	\$0.00
Nov 19	2	0	\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00	\$0.00
Nov 20	15	17	\$1,777.50	1	\$36.95	\$1,814.45	\$3.10	\$336.00	\$2,153.55
Nov 21	17	24	\$2,640.00	4	\$177.00	\$2,817.00	\$14.88	\$313.00	\$3,144.88
Nov 22	10	8	\$945.00	5	\$238.00	\$1,183.00	\$20.01	\$77.50	\$1,280.51
Nov 24	2	0	\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00	\$0.00

Page 1 of 2

Date	# Sales	# Services	Service Sales	# Products	Product Sales	Net Total	Taxes	Tips	Gross Total
Nov 25	3	0	\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00	\$0.00
Nov 26	2	0	\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00	\$0.00
Nov 27	6	0	\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00	\$0.00
Nov 28	16	19	\$2,565.00	2	\$123.00	\$2,708.00	\$10.34	\$490.25	\$3,208.59
Nov 29	10	13	\$1,270.00	0	\$0.00	\$1,270.00	\$0.00	\$284.90	\$1,554.90
Nov 30	22	18	\$1,685.50	12	\$720.80	\$2,406.30	\$60.56	\$293.00	\$2,759.86
Total	242	252	\$25,537.25	65	\$3,924.38	\$29,461.55	\$329.71	\$4,563.23	\$34,354.49

\* The service sales include the value of applied packages.

(c) Worst-scoring head (Head 16)

Figure 7: Attention heatmaps from Layer 15 using three attribution strategies. The top head yields focused and accurate attribution, the mean head shows diluted but somewhat relevant attention, and the worst head highlights largely irrelevant regions.